

The Homophone Effect in Mandarin Word Recognition

Research Thesis

Presented in Partial Fulfillment of the Requirements for graduation

“with Research Distinction” in the undergraduate colleges of

The Ohio State University

By Wei Zhou

The Ohio State University May 2015

Project Advisor:

Professor Shari Speer, Department of Linguistics

Dr. Kiwako Ito, Department of Linguistics

Introduction

Compared to English, Chinese is an ambiguous language. For example, in Chinese, the sentence *ta1 shi4 ge4 mang2 ren2* has two distinct interpretations. One is ‘he is a busy man’, “他是一个忙人.” The other is ‘he is a blind man’, “他是一个盲人.” (Wiener, Speer & Shank, 2012). Without further context or written information, the correct interpretation is impossible to determine. Such lexical ambiguities, unfortunately, are ubiquitous in Mandarin due to the large number of homophones. Mandarin Chinese has 1300 distinct syllable-tone combinations with an average of 5.4 morphemes per syllable (Duanmu, 1999). Figure 1 is the histogram of homophone frequency.

This thesis investigates how native Mandarin speakers use lexical tone information of words that have varying homophone density for word recognition. The study is built on a previous eye-tracking study by Wiener and Ito (in press), which demonstrated that listeners’ knowledge about tonal probability (i.e., the likelihood of tone for the given syllable) facilitates the recognition of spoken words in Mandarin, but only for the infrequent syllables. That is, listeners were faster in recognizing the target character when the stimulus was an infrequent syllable with highly probable tone, whereas they were slower to detect the target upon hearing an infrequent syllable with improbable tone. Such effect of tonal probability was not found for the highly frequent syllables. Wiener and Ito suspect that the lower informativeness of tonal probability for the highly frequent syllables may have to do with their higher homophone density, i.e., tracing the tonal information would not help identify a target word that has many other competitors with the same syllable-tone combination. Since Wiener and Ito did not control the homophone density of the targets, the effect of homophone density could not be examined separately from the effect of syllable frequency. The present thesis project is designed to test

whether homophone density has a consistent effect on the tonal informativeness regardless of the syllable frequency. The following sections provide (1) the summary of Mandarin syllable and word structures that may give rise to a wide range of homophone density; (2) the past studies on the effect of homophones on word recognition; and (3) the design and the results of Wiener and Ito that this project is built on.

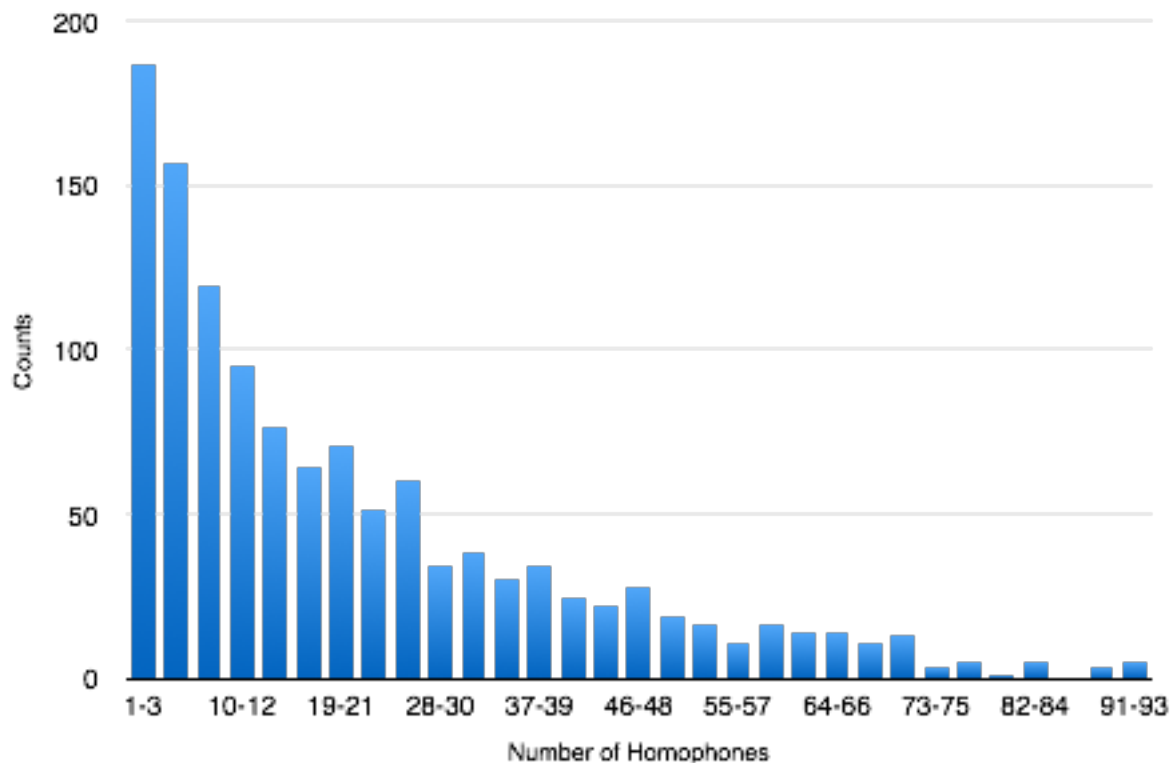


Figure 1: Histogram of Homophone Frequency

1.1 Mandarin syllable/word structures

Two structural characteristics of Mandarin may have to do with the high number of homophones: the syllable structure and the word structure. Chinese has a small number of syllables. The prototypical Chinese syllable is composed of four phonemes, CGVX, where G is a glide and X can be either a vowel or a consonant (Duanmu, 2012). There are 18 consonants that can occur in the syllable-initial position, 3 glides, and 5 single vowels (Duanmu, 2012). Also, four possible phonemes in syllable-ending position (X) are composed of 2 vowels and 2

consonants (Duanmu, 2012). In addition, the onset consonant, glide, and syllable-ending position (X) are optional in some cases, such as 鵝 (*e2*; ‘goose’) and 愛 (*ai2*, ‘love’), and it theoretically increases possible syllables. Hypothetically, 1900 possible syllable should be generated based on the combinations of C, G, V, and X. However, there are only 400 attested (C)(G)V(X) syllable types in Mandarin (Duanmu, 2012) much fewer than the 15000 syllable types in English. As Duanmu (2012) summarizes, some tentative explanations include the conflicting acoustic features of the consonant and vowel such as [yi] where [y] and [i] are both round but have the opposite values.

As to word structure, monosyllabic and disyllabic words are dominantly used in Mandarin. The number of monosyllabic and disyllabic words outweighs that of other types of words combined. Seventy-five percent of total words have less than 2 syllables (Chen et. al., 1993). Also, monosyllabic and disyllabic words are used dominantly in everyday communication. About 96 out of 100 used Mandarin words are either mono- or di-syllabic (Chen et. al., 1993; Jin, 2011). On the other hand, though there is a great portion of di-syllable words, as many researchers suggest, the di-syllabic word structure is a way to reduce homophone size. Ke (2006) proposes that language is a “self-organizing system” that can reduce its ambiguity through the language structure itself or communicative interactions. Chinese has a small number of possible syllable types to form words. The small inventory of syllables restricts the number of distinctive word forms that Mandarin can have and consequently generates numerable pairs of homophones. Disyllabification as one way to increase the syllable inventory by increasing the length of word syllables may reduce the homophone size. As suggested, the negative correlation between the syllable inventory and degree of disyllabification was speculated and found among Chinese dialects (Ke, 2006). In addition, Duanmu (2012) suggests a concept of ‘dual-

vocabulary'; that is, most monosyllabic words (e.g. 学 *xue2* 'study') have disyllabic forms where only one syllable is emphasized. For example, 学习 (*xue2-xi2* 'study'); consists of 学 (*xue2*, 'study') and 习 (*xi2*, 'practice'; Duanmu, 2012). The adding one syllable to the monosyllabic word does not change the meaning; however, it decreases the probability of ambiguity in communication. Although Chinese as a language can manage its ambiguity through disyllabification, it still has a large number of homophones in monosyllabic words.

Another characteristic of Mandarin that may have contributed to a larger number of homophones is its orthography. The majority of Mandarin words have a one-to-one corresponding relationship between words and characters. Unlike English and other languages with alphabetic writing system, Mandarin uses ideographic writing system where symbols and sounds are not tightly mapped to one another. The connections among phonology, semantics, and orthography, though exist, are not as predictive as language like English. Chinese has certain rules to recognize the phonology and semantics of words through orthography. For example, '日' (*ri4* 'sun') originates from the depiction of the round shape of the sun. As a radical, '日' (*ri4* 'sun') refers to the sun and it can contribute to the semantics of words such as in the case of '旱' (*han4* 'drought') where the top part is '日' referring to the sun and the bottom part comes from another word '干' (*gan1*, 'dry'). In addition, the graphic components can contribute the phonetics of the word (Qiu, 2000). For instance, '焊' (*han4* 'to weld') is combined of '火' (*huo3* 'fire') and '旱' (*han4*, 'drought'), where '旱' (*han4*) functions as a sound reference while '火' (*huo3* 'fire') plays the role of semantic indicator for '焊' (*han4* 'to weld'). However, although Chinese words may be recognize by knowing part of Chinese radicals such as the example discussed above, the rule of pronunciation does not work perfectly and exceptions can occur

often such as ‘青’ (*qing1*, color ‘green’, ‘blue’, or ‘dark’) – ‘清’ (*qing1*, ‘clear’) – ‘倩’ (*qian4*, ‘pretty’) – ‘菁’ (*jing1*, ‘luxuriant’). In addition, visually similar words that share parts of their radicals may vary in meaning. In the example above, the common radical is ‘青’ (*qing1*, ‘nature’s color’, ‘green’ or ‘blue’, or ‘young’). Adding ‘氵’, which refers to word ‘水’ (*shui3*, ‘water’), is ‘清’ (*qing1*, ‘clear’) that describes the nature color of water, that is, crystal and clear. Adding ‘亻’, which represents the ‘人’ (*ren2*, ‘person’), is ‘倩’ (*qian4*, ‘pretty’) that depicts the beauty of a young person. Moreover, words that have similar phonology can have different origins and thus distinct meaning words may have distinct characters. For example, words that are pronounced as *yi4* can be written as ‘亿’ (‘hundred million’), ‘易’ (‘easy’), or ‘翼’ (‘wing’). They share the same phonology but contrast in meaning, radicals, and number of strokes, a concept to describe the ‘basic unit of handwriting’ (Qiu, 2000). Summarily, there are some rules that can apply to complex words, but they are not definite. Unlike English, whose alphabetic writing system is based on majority phonology, the writing system of Chinese is mostly based on graphically represented semantics and borrowed phonetic components (Qiu, 2000). Consequently, the homophone may have a more important role in Mandarin recognition.

With such a small number of syllable types and short words, how do native Mandarin speakers understand each other when they communicate without text? One distinctive phonological characteristic of Mandarin is lexical tone – the melody for each syllable that distinguishes word meanings. Mandarin has four lexical tones: high tone (indicated as 1 in pinyin, the alphabetic transcription system of Mandarin words), rising tone (2 in pinyin), dipping tone (3 in pinyin) and falling tone (4 in pinyin). For example, the syllable *tang* [taŋ] can combine with a high, song-like tone (*tang1*, ‘soup’), a rising, question-like tone (*tang2*, ‘sugar’), a low dipping tone (*tang3*, ‘to lie down’), or a falling tone (*tang4*, ‘to scald’). Not all of the 400

syllables can have all four tones, yet once tones are considered, Mandarin has a total of 1300 unique syllable-tone combinations. This may suggest that lexical tone may play a very significant role in spoken word recognition in Mandarin.

Effect of Homophones on Word Recognition

How homophone mates influence word recognition is highly debated. According to Pexman, Lupker, and Jared (2001), the feedback from phonology to orthography was accountable for the inhibitive homophone effect in Lexical Decision Task (LDT). Pseudohomophone non-words, or foils (e.g. *Brane- Brain*) were used in the reported experiments to investigate the possibility of reducing homophone effect as proposed by the authors. In the six reported experiments, both between-subject and within-subject designs were used to investigate the possible of homophone effect in LDT. Word-like visual stimuli were used, including low frequency words with high-frequency homophone mates (low-high homophones), low-frequency words with low-frequency homophone mates (low-low homophones), high-frequency words with low-frequency homophone mates (high-low homophones), and high-frequency words with high-frequency homophone mates (high-high homophones). Participants were presented with foils and instructed to make lexical decisions. Across the experiments, results consistently showed that only low-high homophone group were influenced by the word-like pseudoword foils and more the foils were orthographically similar to real words, larger the homophone effect was. However, the testing of phonology-deemphasizing strategy, which was through presenting word-like foils first and then pseudohomophones, did not reveal significant result. It suggests that phonological feedback activated the orthography of high frequency homophone, therefore resulting in slow responses of low frequency homophones. The phonology-orthography relation was revealed to be the possible cause of homophone effect in LDT (Pexman et. al., 2001).

Several studies of context effect on lexical processing suggest that contextually related homophones may be all activated and competing with each other; meanwhile, the frequent homophone mates are activated more quickly. Onifer and Swinney (1981) designed two experiments to test the influence of frequency and context. In the studies, the frequency-based primary and secondary interpretations of words were presented in different sentences, which might bias to either one of the two interpretations or be ambiguous. A visual word was presented at the end of each sentence and participants were instructed to make lexical decision on it. In the first study, there was no delay between the sentence offset and the presentation of visual words, while in the second study a 1.5s delay (vs. no delay) was added as another factor. The result in study 1 showed that decisions of primary and secondary meaning in both biased and unbiased context were facilitated by the visually presented words. However, besides replicating the results of study 1, results in study 2 showed that only contextually biased meaning was facilitated by the visually presented words in the delay condition. As the Onifer and Swinney (1981) conclude, all meanings of the homophones are activated for at least a short period time before any decision is made.

Similar but different results were found in the context effect in Chinese. Li and Yip (1996) indicate that the Chinese lexical processing in sentence cannot be simply explained by context-dependency hypothesis or exhaustive activation hypothesis; instead, the context effect in Chinese lexical processing needs to incorporate both interpretations: The activated homophones in sentence context are restricted by the context as context-dependency hypothesis suggests, but the lexical access is based on bottom-up processing. In the study, the interpretations of words (primary vs secondary meaning), the nature of sentence (biased or unbiased), probe type, homophone density, and stimulus onset asynchrony (onset vs offset) were manipulated as

independent variables. Participants were instructed to hear a sentence and articulate the word which appeared on the screen, a word that was biased to the context, unbiased, or unrelated. Based on the condition participants were assigned to, the visual word was presented either simultaneously with the auditory homophone at the end of each sentence or after the auditory stimuli. The results showed that contextually biased meaning facilitated the naming, indicating that context restricted activated homophones to semantically related homophones. Also, the interaction of interpretation of meaning, probe type, and homophones indicates that frequency influenced the response speed of words with fewer homophones in unbiased condition. However, when the sentence was biased, there was no homophone effect. In other words, homophone effect occurs only when the context did not provide additional information for lexical access (Li & Yip, 1996).

Caramazza, Costa, and Miozzo (2001) propose a different view that homophones have no effect on word recognition and it is a cross-language phenomenon. In the article, Caramazza and colleague (2001) argue for independent representation (vs. shared representation) of homophones. Independent representation models suggest a single-layer lexical node, that is, lexeme node. Homophone pairs are independently stored and represented mentally. Consequently, the semantic activation of low-frequency words should not be influenced by the high-frequency homophone mates. The speed of lexical processing of low-frequency words should not be facilitated by their high-frequency homophones. Accordingly, the hypothesis suggests that specific frequency of words should be the predictor of recognition speed, rather than the cumulative frequency of all homophonic words. In the study, picture-naming tasks were designed and three groups of stimuli were prepared: A homophone name group, a group of words that matched specific frequency in homophone name group, and a group of words that

matched cumulative frequency in homophone name group. The results in English showed that the response times were more similar between words matched with specific frequency, not cumulative frequency. It suggests that homophone had no effect on word recognition. Reported replications in Chinese and Spanish supported such a claim (Caramazza et. al., 2001).

Contrary to Caramazza et al. (2001), a case study of German aphasia patient showed the opposite view that shared representation of homophones is possible. Contrasting to independent representation hypothesis, shared representation hypothesis suggests that the lexical level can be separated in to two sublevels – lemma and lexeme nodes (Caramazza et. al., 2001). In lemma nodes, grammatical information is stored whereas the phonetic information is stored in lexeme level. According to shared representation hypothesis, homophones are connected through common lexemes. Thus, the activation of one of the homophones mates should affect the activation of other mates through the shared lexeme. Biedermann, Blanken, and Nickels (2002) designed a study to test the association between homophones. In the case study, an aphasia patient experienced a 10-day therapy session on vocabulary. During the therapy, the patient was trained with some German words. The trained words were further divided into four groups based on their paired counterparts – homophones (e.g. trained *ball* vs. untrained *ball*), phonologically related (e.g. trained *vase* vs. untrained *nose*), semantically related (e.g. trained *window* vs. untrained *door*), and unrelated words (e.g. train *violin* vs. untrained *apple*). Three auditory lexical decision tests – a pre-test, a posttest1 (one day after the therapy), and a posttest2 (a week after the therapy) – were given to evaluate the effect of therapy. In the test, both the treated and untreated words were tested. The results showed that only the performance of untreated homophones showed an improvement and no other untreated words did. The authors conclude

that the results supported the one phonological representation for both homophones and hence the shared representation hypothesis (Biedermann, Blanken & Nickels, 2002).

A study on Chinese visual word recognition also suggests a shared phonological representation. Contrary to the homophone study in English, Chen Vaid, and Wu (2009) suggests that the lexical access of low visual frequency words can be facilitated by homophone density. In the study, a LDT and a naming task were conducted to test their hypothesis. In both tasks, identical Chinese compounds were used and categorized based on visual frequency, homophone density, and phonological frequency. In LDT, the result revealed advantages for high visual frequency words and high homophone density words. There was no significant effect of phonological frequency or significant interaction. In naming task, advantages for high visual frequency and high homophone density were showed again. In addition, high phonological frequency items were responded more slowly and less accurately. Combining the results of the two studies, Chen et al. (2009) conclude that the large number of homophones facilitates lexical access. .

Interestingly, the different homophone effects were resulted from different types of tasks and different languages that were studied in research. The inhibitive homophone effect was found in LDT by using pseudo-word foils and it showed only among low frequency words with high homophone mates. Studies of context effect in Chinese and English showed different mechanisms of lexical processing of homophones. The independent representation of homophones was tested by picture naming tasks and confirmed by measuring and comparing response time. The shared representation of homophones was supported by the accuracy measure of an auditory LDT and in the study of Chinese compounds homophone density was manipulated as an influential factor. Therefore, a question rises: Are the used tasks or tested language

influencing the mental processing? The answer is yes and some study demonstrated the influences of experiment design on mental processing.

According to Kerswell et al. (2007), lexical decision task (LDT) requires integrating more information from orthographic representation, but semantic categorization task (SCT) requires incorporating more information at semantic level. In the study, they compared two groups of homophones pairs – words with high frequent homophone mates that are not animal names and words with low frequency homophone mates that are animal names – in both LDT and SCT. What they found is that nonanimal homophone mates were responded more slowly than control in LDT, but similar to control in SCT, while animal homophone mates were responded less slowly than control in LDT, comparing to the large difference in SCT. It suggests that in LDT, the feedback from phonology to orthography overweighed the feedback from phonology to semantics and competition between different orthographies obstructed the lexical access, resulting in slower responses to non-animal homophone mates with high frequency; in SCT, the feedback from phonology to semantics was more dominant such that the semantic competition between homophone mates inhibited the lexical access, consequently generating slower responses for low frequency homophone mates (Kerswell et. al., 2007).

Similarly, one study about spelling indicates that homophone effects can either be facilitative or inhibitive, depending on the types of tasks. Triman, Seidenberg and Kessler (2014) assessed participants' preference on spelling of novel make-up homophones. Triman and colleague made up some novel words and either asked participants to spell out the homophones themselves, or provided them with two options including a novel spelling based on correct phonology rules and a familiar spelling. Even when the metalinguistic focus was increased through instructing participants to explain why they choose the spelling, familiar spellings were

more preferred by those who were instructed to create the new spellings. However, when participants were provided with two options, the unfamiliar spellings were preferred. The author concludes that different pressures motivated people to use different strategies regarding to tasks at hand. It suggests that depending on the tasks, the knowledge of stored homophones can facilitate learning, but the competition among spellings can also be a problem to differentiate the meaning of homophones (Triman, et. al., 2014).

Summarily, previous literature suggests that homophones influence lexical processing in word recognition, competing or facilitating each other. Frequency as an additional factor contributes to the mental processing of lexicons. Importantly, most of literature suggests the connections among homophones and the influence of homophones. Therefore, the homophone density convincingly influences word recognition.

1.3 Review of Wiener and Ito (in press)

According to Wiener and Ito (in press), Chinese tonal information can be useful in word-recognition. Specifically, what they found is that tonal probability was useful in identifying low-frequency words but there was no effect on lexical access for high-frequency words. In the study, clicking responses were recorded. An eye-tracking device was used to help to interpret the unconscious lexical processing of Mandarin Chinese. Twenty-four pairs of mono-syllabic Mandarin words that contrasted in tonal probability were selected as experiment items, in a total of 48 items. Half of the items had low syllable frequency, whereas the other half had high syllable frequency. Participants who were mono-dialectal Mandarin speakers were instructed to listen to the auditory stimuli of experimental items and select the one that they heard from the four options presented on computer screen: target, segmental competitor, rhyme competitor, and distractor. Eye movement and response time were recorded. The analysis of fixation observations

showed a clear competition between the tonally contrasted pairs of words. The click responses showed that items with high (vs low) tonal probability were responded faster only in low syllable frequency condition, while in high syllable frequency conditions, the difference between high and low tonal probability items was not significant. The results suggest a dissymmetrical use of tonal information in spoken word recognition.

However, as suggested by Wiener and Ito (in press) in the discussion, the large number of homophones may potentially be a problem for the null difference of tonal probability in high syllable frequency condition. For example, they pointed out that the number of homophones of each item in high syllable frequency conditions was above 20. The distinction of high vs low tonal probability was not able to effectively reduce the size of homophones among frequent syllables and the large homophone number might significantly inhibit the lexical access processing (Wiener & Ito, in press). In addition, another question that may rise from the uncontrolled homophone density is that items in low syllable frequency and high tonal probability condition may have more homophones than those in low syllable frequency and low tonal probability, therefore resulting in faster responses. Both items in high and low tonal probability in low syllable conditions may have fewer homophones than items in high syllable conditions. However, probable tone words in low syllable frequency words may have the higher number of homophones did improbable tone words because items with more homophones may indicate a larger searching area in memory and thus higher likelihood of being picked out, compared to those with fewer homophones.

Motivated by such a perspective, the current study further investigates the function of tonal probability in Mandarin spoken word recognition. Following the experimental design that was used in Wiener and Ito (in press), the current study continues recording click responses and

unconscious eye movement to help to understand the mental processing of lexicons. Besides the two factors controlled in Wiener and Ito, an additional factor – homophone density – was added as independent variable. Based on the results of Wiener and Ito (in press) and previous literature on homophone effect, the following hypotheses are speculated:

- a) Items with high homophone density, low syllable frequency, and low tonal probability will be the slowest among all conditions.

As showed in Wiener and Ito (in press), low syllable frequency and low tonal probability generated the slowest responses among all four conditions. If high homophone density has an inhibitive effect on word recognition, adding another inhibitive factor beside low syllable frequency and low tonal probability will result in slowest responses.

- b) Items with low homophone density, high syllable frequency, and high tonal probability will be the fastest among all condition.

As suggested by Wiener and Ito (in press), if the uncontrolled high homophone density might be a great influence on response time in high frequency conditions, controlling the confounding homophone density will substantially reduce the competition from other frequent items with probable tone. In addition, items with high frequency has showed to be more accessible in studies of Chinese (Caramazza et. al., 2001) and high tonal probability has showed the advantage of being recognized faster in the low syllable condition (Wiener & Ito, in press). It is reasonable to believe that participants will respond most quickly to items with low homophone density, high syllable frequency, and high tonal probability.

- c) The following three conditions will form the second fastest group: High homophone density, high syllable frequency, and high tonal probability condition, low homophone density, high syllable frequency, and high tonal probability condition, and low homophone density, low syllable frequency, and low tonal probability condition.

low homophone density, low syllable frequency, and high tonal probability condition, whereas these three conditions will form the second slowest group: high homophone density, high syllable frequency, and low tonal probability condition, high homophone density, low syllable frequency, and high tonal probability condition, and low homophone density, low syllable frequency, and low tonal probability condition. Because high homophone density, low syllable frequency, and low tonal probability are considered to be inhibitive, whereas low homophone density, high syllable frequency, and high tonal probability are considered to be facilitative, the combination of two facilitative factors along with one inhibitive factor is likely to generate the second fastest responses, and the combination of two inhibitive factors along with one facilitative factor is likely to generate the second slowest responses.

Methodology

Participants

Forty-nine Mandarin speakers from China, who were students at OSU, were recruited for a small incentive. Two participants did not complete their sessions due to technical problems and were excluded from further analysis. The average age of the remaining 47 participants (19 males, 28 females) was 21.85 years old. All participants obtained at least high school degree in China before coming to the United States. Twenty students attended college or obtained their college education in China before studying at OSU. Only one participant had attended a graduate school in China. The mean length of their staying in United States by the time they completed their session was 21.5 months and it ranged from a few days to 53 months.

Design and Materials

Three factors – homophone density (high vs low homophone: H+, H-), syllable frequency (high vs low frequency: F+, F-), and tonal probability (high vs low probability: P+, P-) were manipulated. To determine homophone density, items that have 7 or more homophones were categorized as having a high homophone density while items that have between 2 and 6 homophones were categorized as having a low in homophone density. The chosen cut-off of 7 was due to the practical reason that we were not able to find enough experimental items for each condition if it was a larger number. To be considered as a high-frequency syllable, the syllable-log frequency needs to be 4.8 per million or larger, while to be considered as a low-frequency syllable, the syllable-log frequency needs to be between 1.0 and 4.0 per million. The tonal probability was categorized based on the tone-log frequency, where the high probable tone was the one that has the highest tone-log frequency and the low probable tone that has the lowest tone-log frequency. All critical items had a syllable that can be pronounced with 3 or 4 tones.

		Homophone Density			
		High		Low	
Syllable Frequency		High	Low	High	Low
Tonal Probability	High	H+F+P+ e.g. 枷, <i>jia1</i> 'scaffold'	H+F-P+ e.g. 频, <i>pin2</i> 'frequency'	H-F+P+ e.g. 孩, <i>hai2</i> 'child'	H-F-P+ e.g. 狂, <i>kuang2</i> 'mad'
	Low	H+F+P- e.g. 贺, <i>he4</i> 'to congratulate'	H+F-P- e.g. 掺, <i>chan1</i> 'to adulterate'	H-F+P- e.g. 详, <i>xiang2</i> 'detailed'	H-F-P- e.g. 俗, <i>su2</i> 'common'

Table 1: Eight Conditions and a Sample Item within Each Condition

The three factors were crossed to create 8 conditions in a 2 x 2 x 2 design. Table 1 shows the 8 conditions with a sample item in each condition. A total of 96 items were selected. There were 32 items, four in each experimental condition, and 64 fillers. Two lists were created, each

including all 96 items. In each list, items were pseudo-randomized such that an equal number of items from each condition appeared in the first and second halves of the list.

The order of the two halves in the list was counterbalanced such that the two halves were presented both in the first and the second halves cross sessions. Half of participants received one version and the rest participants received the other.

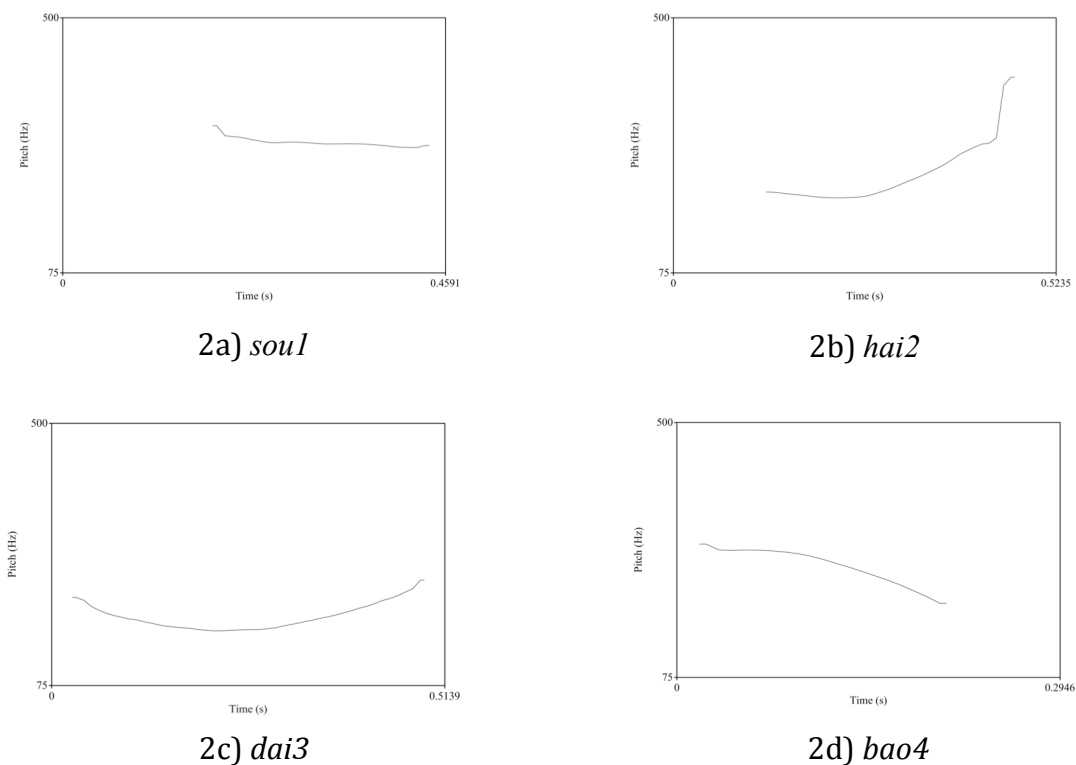


Figure 2: Samples of f0 contour of Mandarin tones

Auditory stimuli

The auditory stimuli were recorded by a 28-year-old Chinese woman from Beijing. In all experimental conditions except for H+F-P- condition, all four Mandarin tones were represented once. In H+F-P- condition, tone 1 word was used twice because of the lack of a candidate with tone 4. The samples of pitch contour for four tones were presented in Figure 2: 2a. *sou1*, 2b. *hai2*, 2c. *dai3*, 2d. *bao4*.

Table 2 shows the mean word duration for each condition. For the filler trials, a total of sixty-four words were selected and recorded. The most selected tone of fillers was tone 4, then tone 1, tone 2, and lastly tone 3, such that the tonal distribution of fillers was consistent with the natural distribution in Mandarin. Each target word was presented after a carrier phrase followed by a pause: *wo3 yao4 shou1*, “I am going to say,”.

Condition	Mean Word Duration (ms)
H+F+P+	421
H+F+P-	510
H+F-P+	479
H+F-P-	469
H-F+P+	512
H-F+P-	489
H-F-P+	527
H-F-P-	482

Table 2: Mean Duration of Auditory Words in Each Condition

Visual stimuli

A total of 96 slides (32 critical, 64 fillers) were created. In each slide of a critical trial, four different monosyllabic characters: Target, segmental competitor, rhyme competitor, and distractor, were presented. A segmental competitor shared the syllable with the target, but contrasted in tonal probability. For example, if the target had the most probable tone, then the competitor had the least probable tone, or vice versa. A rhyme competitor had a different onset consonant but the same rhyme as the target. A distractor was acoustically distinct from all other words. Figure 3 shows a sample slide of a critical trial in H-F-P+ condition. The target (*zhai4*, ‘debts’: bottom right) had low homophone density, low syllable frequency, and high tonal probability. The segmental competitor (*zhai3*, “narrow”: bottom left) shared the same syllable but had the least probable tone. The rhyme competitor (e.g. *tai4* ‘to discard’: upper right) had the identical vowel and tone with the target. The distractor (e.g. *qun2* ‘skirt’: upper left) was

completely different from other three words. The character-log frequency of each character was above 1.0. Five mono-dialectal college students in China confirmed that they easily recognized all critical items and their segmental competitors. Also, the character frequency and strokes of each character within a slide were approximately matched as well. None of the simultaneously presented characters shared any radical. However, we could not control the homophone density of non-target words (i.e., segmental competitor, rhyme competitor, and distractor).

Sixty-four fillers included 32 semantic and 32 visual fillers. In each semantic filler trial, there was one character that was semantically related the target (e.g. target 罚 *fa2* ‘to punish’;

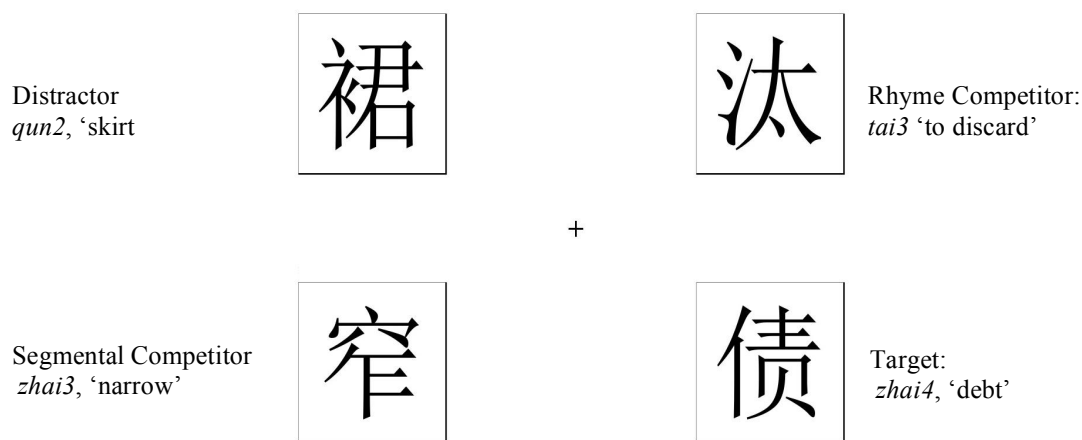


Figure 3: An Example Slide of Critical Trial in H-F-P+ Condition

semantically related competitor 赏 *shang3* ‘to reward’). In each visual filler trial, there was one character that looked similar to the target (e.g. target 句 *ju4* ‘sentence’; visual competitor 勾 *gou1* ‘to evoke’). Two distractors with different syllables and tones than the target filled the remaining two positions in each of the filler slide.

Procedure

Participants first completed a questionnaire about their language background, questions such as ‘how long have you been in the U.S.’, ‘what other Chinese dialects do you speak?’, and

‘on average, how many minutes a day do you spend reading or writing in Chinese (See Appendix for a translated version). After completing the questionnaire, they were calibrated with ClearView 5-point calibration program on Tobii 1750 monitor. The experiment started with a brief auditory instruction in Mandarin. A total of 96 trials (32 criticals, 64 fillers) were presented in a pseudo-randomized order. In each trial, participants were instructed to fixate upon the cross at the center of a blank screen for about 2 seconds. Then, they heard the carrier phrase, *wo3 yao4 shuo1*, ‘I am going to say’, and saw a slide with four characters presented together including the target word. Participants were instructed to search freely and to click on the word they just heard. The fixation data were collected continuously by Tobii 1750 system at 50Hz sampling rate. After finishing their session, they were debriefed and thanked, and received a small incentive.

Result

Responses from 47 participants were analyzed after removing data from two participants who did not finish their sessions.

Click response. Participants’ response times were calculated as the lag between offset of target word and the time of clicking. Sixty-three errors and missed responses were excluded from further analysis. Twenty-two outliers that were more than 3 standard deviations away from the participant’s mean within each condition were excluded. There were 1418 remaining observations across the 8 conditions. Figure 4 shows the average response times for the eight conditions.

Table 3 summarizes the outcome of the mixed effects models using lme4 package of R 3.0.2 (Bates, Machler, Bolker, & Walker, 2014). The model included homophone density, frequency, tonal probability and their interactions as fixed effects, subjects and items as random intercepts, and maximum random slopes (the three within-subject factors and their interactions) for subjects. The *p*-value was calculated based on the method described in Hakekoh & Højsgaard

(2015), using the R package “pbkrtest”. The analysis revealed no main effect, but a significant 3-way interaction (Table 3).

The 3-way interaction was qualified by the only significant difference between H+F-P- and H-F+P+. As expected, H+F-P- was the fastest, whereas H-F+P+ was the slowest. Although none of the main effects produced significant effects, a significant difference between the fastest and the slowest supported the hypothesized effects that when the three factors influenced recognition in the same direction, they generated the fastest and slowest responses. In addition, visual inspection (Figure 3: ‘prob’ – tonal probability; ‘homo’ – homophone density; ‘freq’ – syllable frequency) suggests homophone density might have an effect. As it shows, H+F+P+ was slower than H-F+P+ and H+F-P- was slower H-F-P-. However, current data did not provide statistically evidence to support such a comparison and we are recruiting more participants to tease apart the three factors.

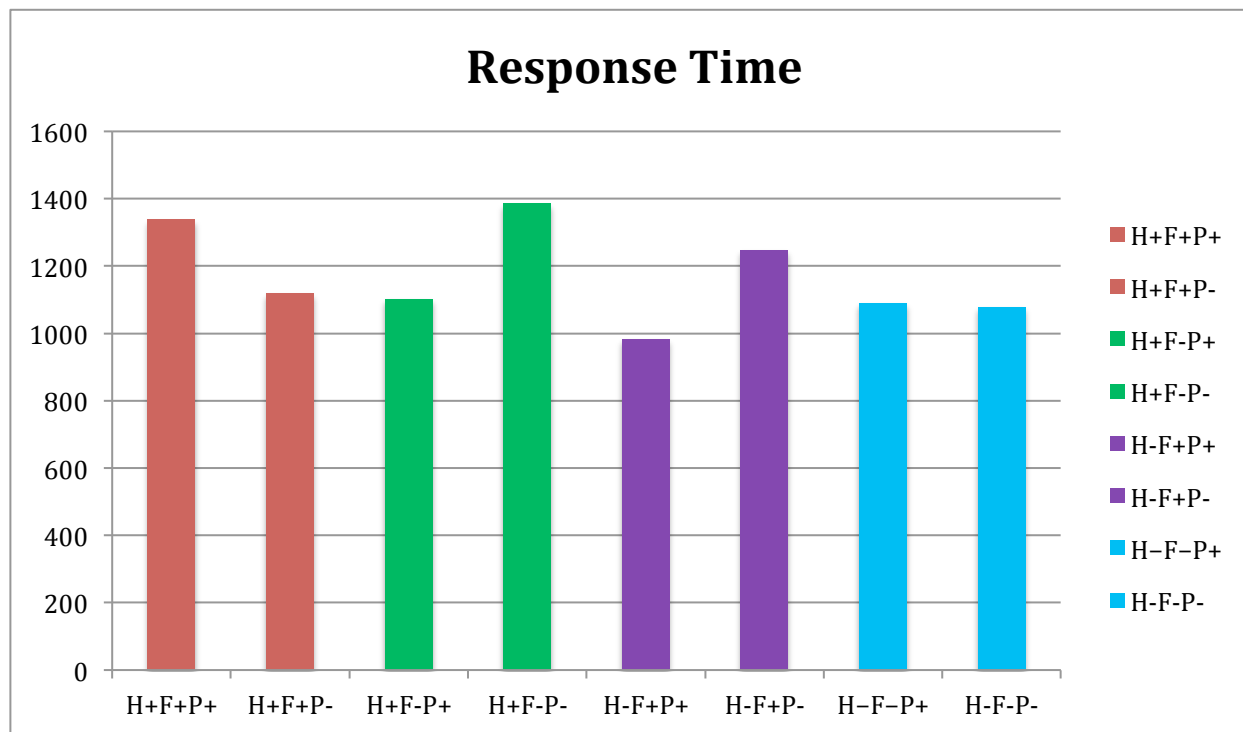


Figure 4: Response Times For 8 Conditions

Fixation patterns. In order to grasp the overall timing of fixations across the eight conditions, the grand-mean functions for the target, segmental competitor, rhyme competitor, and the distractor were plotted (Figure 5). As expected, the segmental competitor that shared the syllable with the target showed the overall highest competition as compared to the rhyme competitor and the distractor. Therefore, the following data analyses focus on a comparison of the fixations on the target and the segmental competitor.

The eye-tracking data from 47 participants showed visible differences in fixation timings across the 8 conditions. Figure 6 shows the time course of fixations to the target and the segmental competitor for each of the eight conditions.

Fixed effects:

	Estimates	Std.Error	<i>t</i> value	<i>p</i> value
(Intercept)	1181.53	58.86	20.073	
prob	-82.19	86.12	-0.954	
homo	135.4	86.42	1.567	
freq	10.6	86	0.123	
prob:homo	55.69	172.08	0.324	
prob:freq	133.15	172.59	0.771	
homo:freq	-79.9	172.65	-0.463	
prob:homo:freq	810.39	343.31	2.361	<.05

Table 3: Analysis Summary of Response Time Mixed-Effects Regression Model

As expected, H-F+P+ showed recognition of target occurred faster than in the other conditions (Figure 6e). The divergent point was around 200ms after the onset and before the offset of stimuli. The target function increased quickly and peaked around 900ms after the onset. The maximum looking proportion also suggested a fast response; around sixty percent of looking was at the target words. On the other hand, the competitor function declined after the divergent point. There was no sign of the competitor competing with the target word during lexical processing.

Consistent with clicking responses, the fixation data suggests that H+F-P- was the slowest (Figure 6d). The target function did not rise until 800ms after onset whereas the average stimuli duration was around 500ms. The maximum proportion of looking at the target decreased greatly comparing to the fastest condition (H-F+P+). Again, the lack of looks to the competitor suggests that it did not compete with target. The competitor function started declining around the offset.

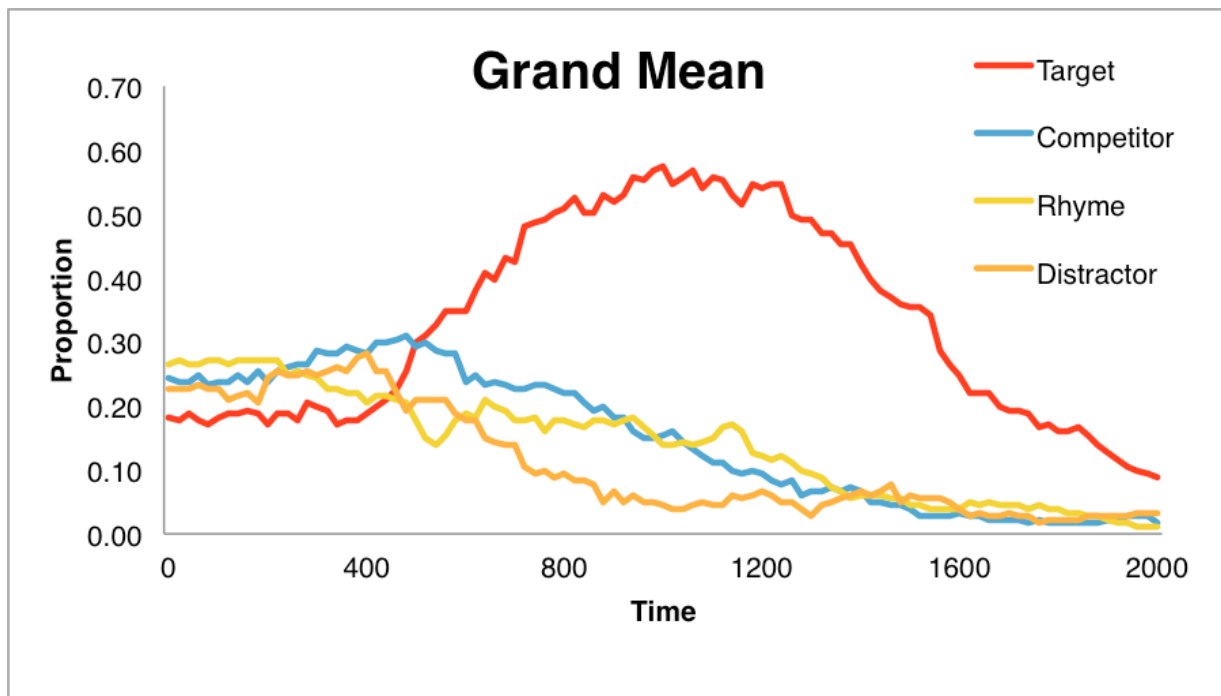


Figure 5: Grand-Mean Functions for Target, Segmental Competitor, Rhyme Competitor, and Distractor

In order to directly compare the fixation timings for the target and for the competitor across the conditions, Figure 7 shows superimposed functions for the target (7a) and for the competitor (7b). We speculated H-F+P+ would be the fastest condition and H+F-P- would be the slowest condition. Consistent with our predictions, the target function (7a) of H-F+P+ (solid purple line) was on top of all other functions, indicating fastest recognition and target function of H+F-P- (dash green line) was the slowest to increase and had one of the lowest maximum

proportions of looking, indicating slowest recognition. The competitor functions (7b) did not show clear trends.

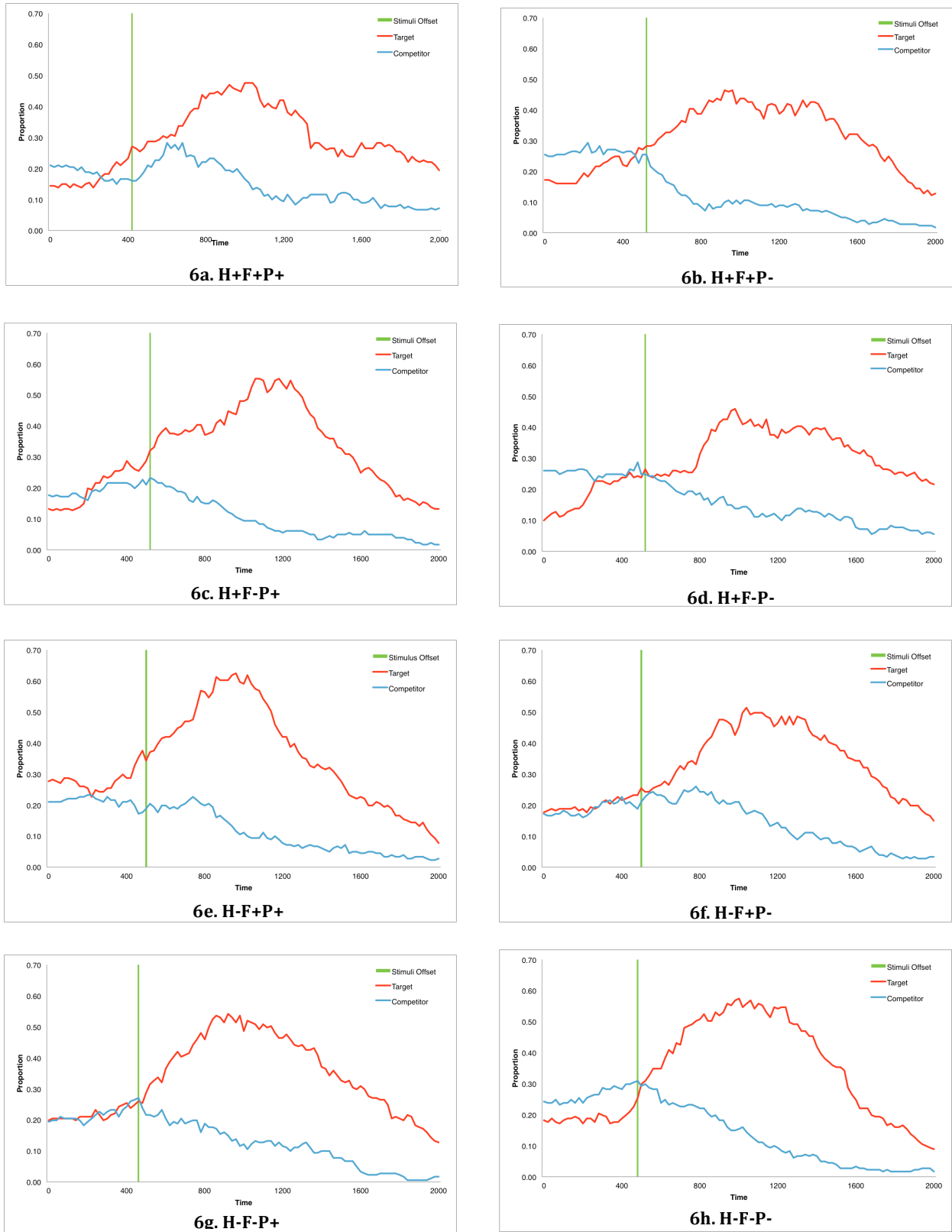
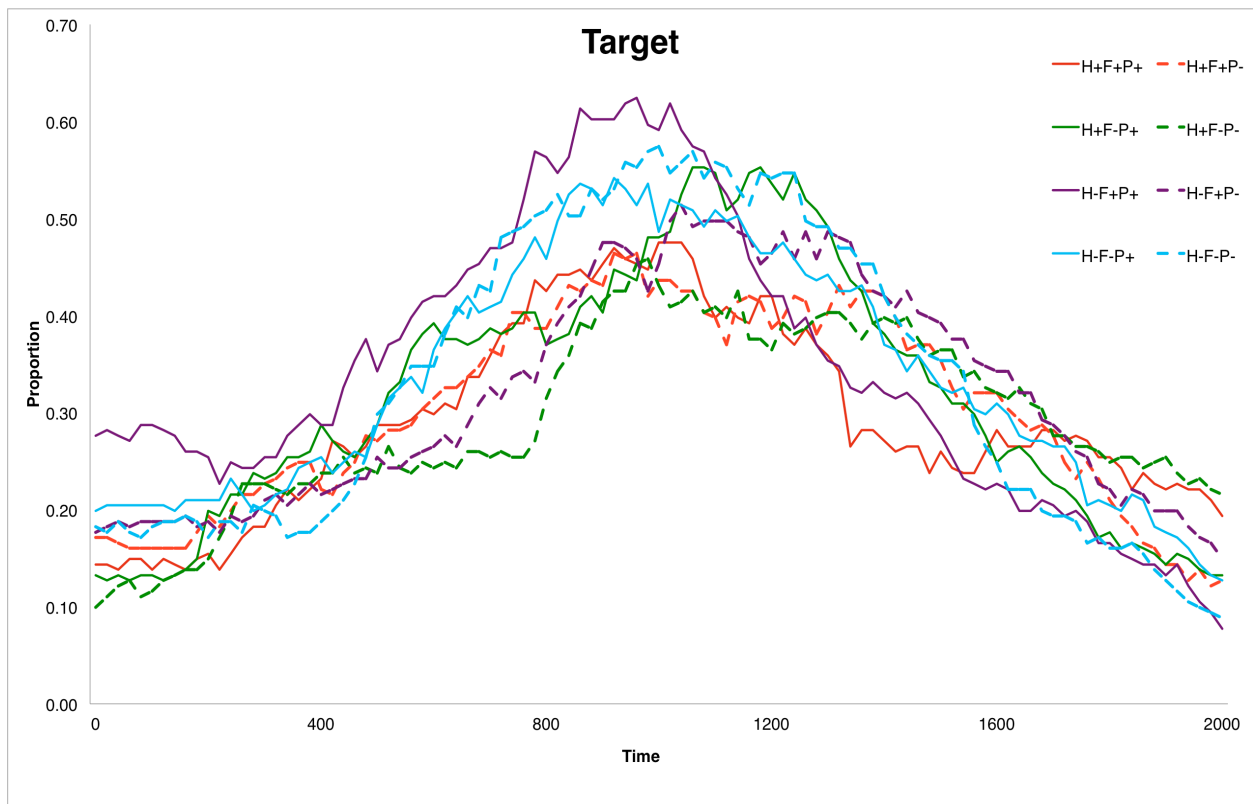
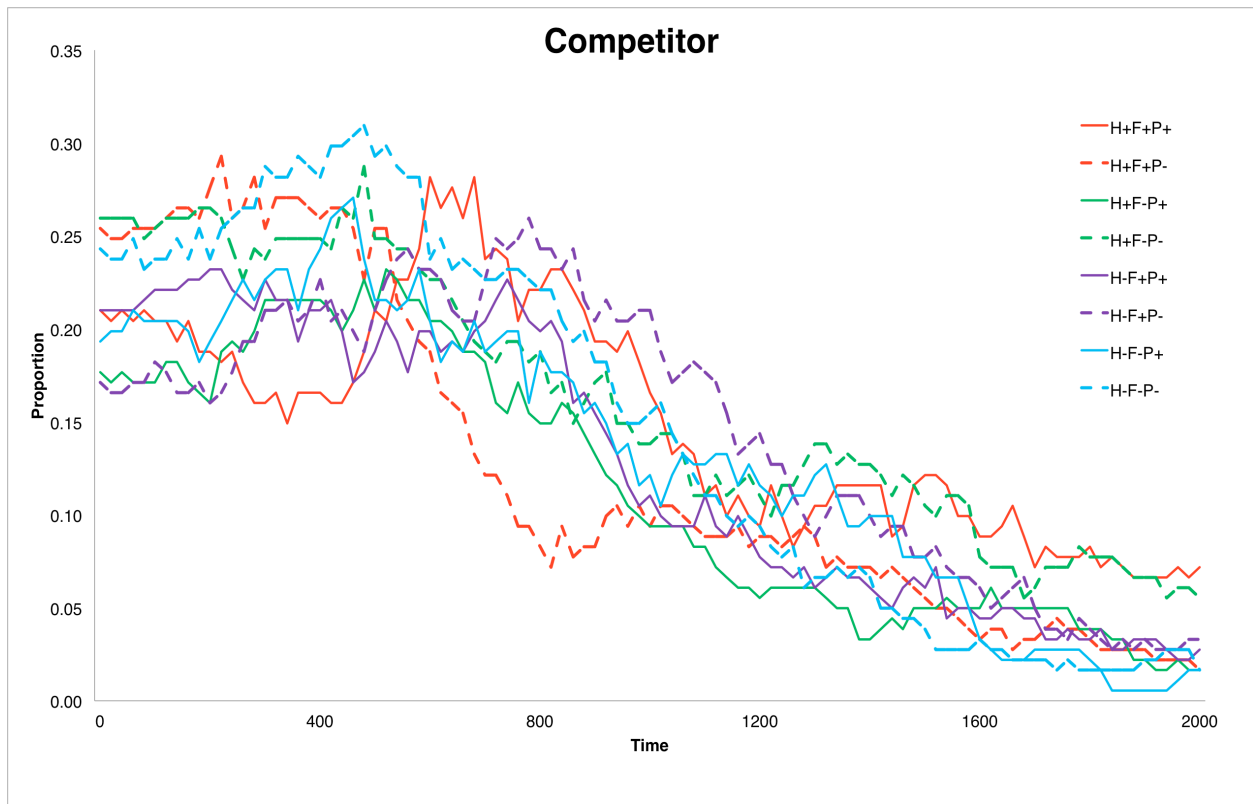


Figure 6: the Time Course of Fixations to the Target and the Segmental Competitor for Each of the Eight Conditions



7a) Target Functions Cross 8 Conditions



7b) Competitor Functions Cross 8 Conditions

Figure 7: Target and Competitor Functions Cross 8 Conditions

TARGET Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.37659	0.03088	-12.197
prob	0.08734	0.05279	1.654
homo	-0.09229	0.0531	-1.738
freq	-0.01093	0.05307	-0.206
prob:homo	-0.06533	0.10522	-0.621
prob:freq	0.0304	0.10525	0.289
homo:freq	0.06637	0.10572	0.628
prob:homo:freq	-0.29474	0.21028	-1.402

COMPETITOR Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.888891	0.079808	-11.138
prob	-0.008277	0.158484	-0.052
homo	-0.046183	0.158578	-0.291
freq	-0.104656	0.158533	-0.66
prob:homo	0.174461	0.316899	0.551
prob:freq	0.246265	0.316901	0.777
homo:freq	-0.013545	0.316901	-0.043
prob:homo:freq	0.271161	0.634148	0.428

Table 4: Analyses Summary of Fixation Mixed-Effect Models

Although the fixation data showed the above interesting trends, the effects of homophone density, syllable frequency, tonal probability and their interactions were not reliable. Based on the grand mean function (Figure 5), which showed the shift of fixations to four candidates starting at around 200ms after the onset of target word and the peak of the fixation rise for the target around 1000ms, the window for the statistical analysis was determined as 200-1000ms from the onset of target. For this time window, the counts of looks to target and to the competitor were converted respectively into empirical logit: $\log(\text{the counts of the looks on the area of interest} / \text{the total observations excluding track loss})$. The mixed effect models for fixation, which had the same specifications of predictor effects, random intercepts and random slopes as the models for the fixation data, revealed no significant effect (Table 4: ‘prob’ – tonal

probability; ‘homo’ – homophone density; ‘freq’ – syllable frequency). We suspect that the lack of reliable effects may be due to small number of observations. Additional participants are currently being recruited to test for possible effect of tonal probability, homophone density, and any three-way interaction suggested by these trends.

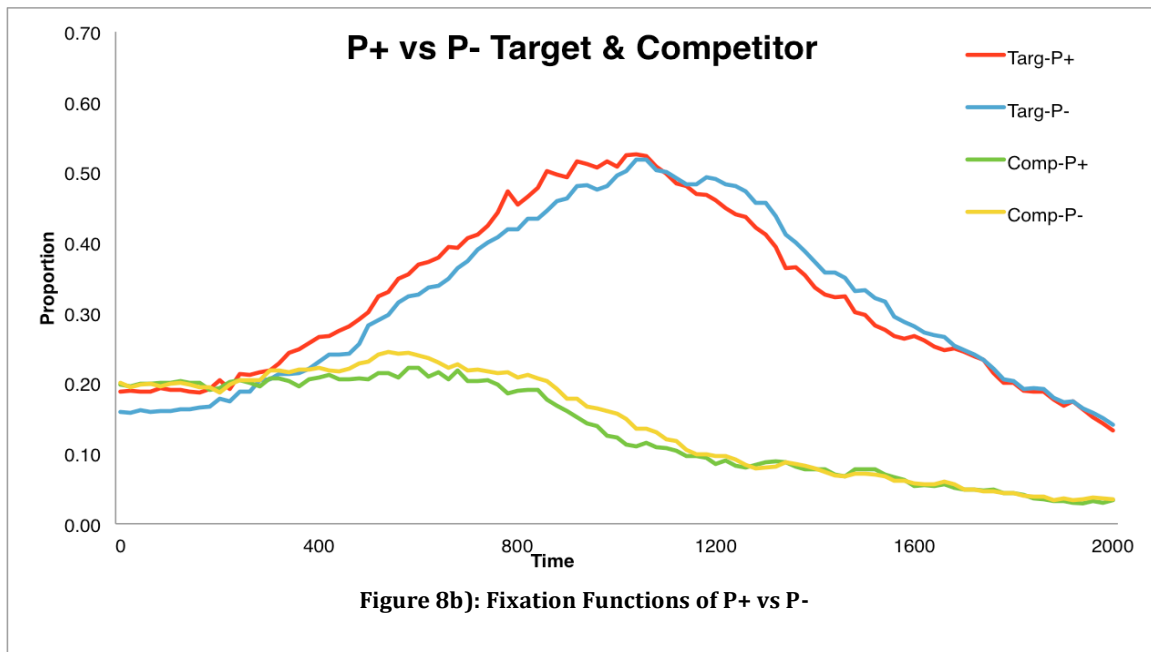
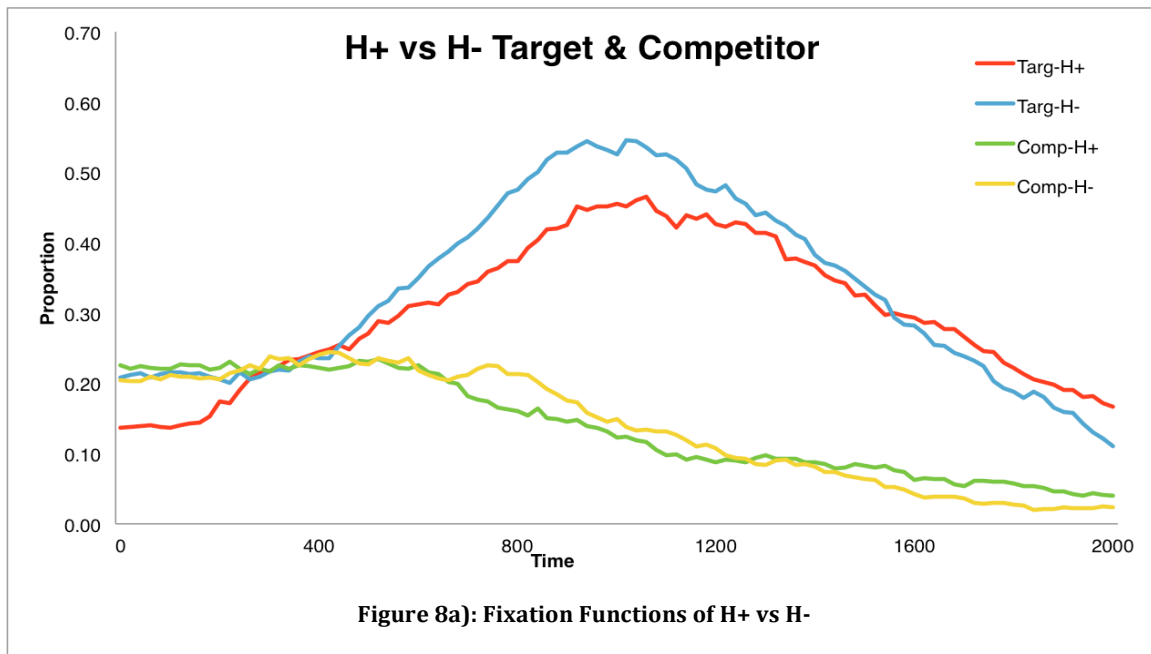


Figure 8: Fixation Functions of Main Effect

Overall Trend

Taken together, the response time data and the fixation data showed numerical trends for the homophone density and the tonal probability manipulations (Figure 8). First, high homophone density tended to slow down the identification of the target. The looks to the target (Figure 8a) were numerically lower for the combined H+ condition (about 46%) than H- conditions (about 54%). The mean response time was 1231.88ms for H+ items and 1089.33ms for H- items, i.e., H+ items were on average 133.55ms slower than H- items. As for the tonal probability, high-probable tones tended to facilitate the target detection. When the target had a probable tone, fixations to the target increased faster, whereas an improbable tone slowed down the fixations to the target, shifting the function to the right (Figure 8b). In addition, participants tended to look more to the competitor with a probable tone when they heard the target with improbable tone. The response times were on average 79.48ms faster for the conditions with probable tones than for those with improbable tones.

Echoing Wiener and Ito, the present data did not show a consistent trend for the overall syllable frequency, i.e., highly frequent syllables were not necessarily processed faster. Homophone density showed numerical interactions with syllable frequency. In high homophone density conditions, frequent syllable indicates a larger number of frequent competitors. Tonal probability in H+F+P+ conditions was not able to effectively reduce the cohort size, as suggested in Wiener and Ito (in press). For example, 懲 (*cheng2*, ‘to punish’; H+F+P+) has 23 homophones including some very commonly used homophone mates such as 成 (*cheng2*, ‘to succeed’) and 城 (*cheng2*, ‘city’ or ‘town’). Competition was increased by the large number of homophones in high frequency condition. However, in low homophone density conditions (H-F+P+), only a few frequently used homophones were available to compete with the target so the probable tone

facilitated the processing. The advantage of frequent syllable then appeared, as H-F+P+ was the fastest condition. Comparing the response time of H+F+P+ vs. H-F+P+, the homophone density effect was suggested as the numerical difference indicated.

The fixation data and response times also echoed each other in showing differences across conditions: participants were fastest to both fixate and click on the target in the H-F+P+ condition, and slowest in the H+F-P- condition. The fixation data for H-F+P+ showed an early divergence between the looks to the target and the looks to the competitor, at around 200ms from the onset. It was around 290ms earlier than the offset of stimuli. That is, participants were able to easily access lexical information pertaining to the targets only after they heard initial part of the word. For example, 孩 (*hai2*, ‘child’: 3 homophones, syllable frequency: 5.43, high tonal probability) was very easily recognized and they did not consider the competitor 害 (*hai4*) much. The fact that the target detection was the fastest for one of the conditions with frequent syllables indicates that, contrary to Wiener and Ito’s findings, the overall segmental frequency may not determine the informativeness of tone.

Interestingly, both the fixation and the response time data showed the slowest detection of the target for the condition that had the opposite-to-the-fastest combination of the three factors: H+F-P-. In this condition, the divergence between the target and the competitor fixations was delayed till the offset of the stimuli. In fact, the looks to the target did not start increasing until about 230ms after the offset of auditory stimuli. That is, listeners identified the target character after they heard the entire word such as 掺 (*chan1*, ‘to adulterate’: 7 homophones, syllable frequency: 3.81, and improbable tone) in the presence of the competitor 钲 (*chan3*). Unlike Wiener and Ito (in press), this condition that led to the slowest detection of the target did not show the highest looks to the segmental competitor.

Discussion

The current study found that a significant three-way interaction that was most apparent from by the significant difference between the fastest (H-F+P+) conditions. This is consistent with the slowest (H+F-P-) suggests predicted effects for our 3 main factors when the homophone density, syllable frequency, and tonal probability drove word recognition in the same direction. The fastest condition (H-F+P+) suggests that segmental frequency did not determine the informativeness of tone. Comparing to the results in Wiener and Ito (in press), the current study suggests a complex interaction of homophone density, syllable frequency, and tonal probability.

What Wiener and Ito found was that segmental frequency determined the informativeness of tone: For high-frequency syllables, there was no effect of tonal probability, but for low-frequency syllables, probable (vs. improbable) tone facilitated recognition. If the uncontrolled homophone density caused the null difference between probable and improbable tone words in high syllable frequency conditions in Wiener and Ito, we would predicted that H+F+P+ trials would be delayed as compared to H-F+P+. In current study, results numerically supported such a speculation in that the response time of H+F+P+ was appeared to be longer than H-F+P+, suggesting the tonal probability was more informative for frequent syllables with low homophone density. The frequent homophone mates compete with each other mentally in H+F+P+. For example, 惩 (*cheng2*, ‘to punish’; H+F+P+) has 23 homophones including *cheng2* (成, ‘to succeed’; 城, ‘town’; 诚, ‘honest’). Comparing to those in H+F+P+, words in H-F+P+ has less competing homophones. For example, 添 (*tian1*, ‘to increase’; H-F+P+) has only one or two frequently used and easily accessible words such as 天 (*tian1*, ‘day’), which facilitated the access.

However, in the low probability condition, the same reasoning was not sustained; high homophone density slowed down lexical processing of improbable tone words. If reducing homophone density among frequent syllables can speed lexical processing, then H+F+P- would be slower than H-F+P-. The result nevertheless revealed the opposite trend: H+F+P- was faster than H-F+P-. One possibility is that in H+F+P- condition, the many high-frequency homophones facilitated the processing, but in H-F+P-, due to the low homophone density, there were fewer homophones that could speed the processing. For example, 腐 (*fu3*, ‘to decay’; H+F+P-) has 20 homophones including common words 抚 (*fu3*, ‘to caress’) and 辅 (*fu3*, ‘to assist’); 印 (*yin4*, ‘to print’, H-F+P-) has only 6 homophones most of which are uncommon or proper nouns such as 胤 (*yin4*, ‘progeny’) and 茛 (*yin4*, ‘an organic compound’). The frequent homophone such as 抚 (*fu3*) and 辅 (*fu3*) can facilitate the processing of their homophone mate 腐 (*fu3*), while 印 (*yin4*) is less likely to receive facilitative effect from 胤 (*yin4*) and 茛 (*yin4*). Therefore, the frequency of homophone may also influence the lexical processing.

In the low frequency conditions, Wiener and Ito indicated an effect of tonal probability in advantage of probable tone in that infrequent syllable with probable tone was responded faster than those with improbable tone. One possible factor that can be considered regarding to the results in Wiener and Ito is the homophone density that may be indicated by tonal probability in infrequent-syllable conditions. High tonal probability may indicate more homophones, while low tonal probability may indicate fewer homophones. If the slower responses in F-P- condition in Wiener and Ito indicates poor performance in recognition tasks, then increasing the homophone density (H+F-P-) should further inhibit the processing therefore having slower clicking responses. The results in current study suggest that the lexical processing of infrequent syllable with improbable tone was inhibited by high homophone density: the numerical value of response

times of H+F-P- and H-F-P- showed such a relation that H-F-P- was faster than H+F-P-. On the other hand, if infrequent syllable items with high probable tone (F-P+) had a larger homophone size than those with improbable tone (F-P-), then increasing the homophone density would reduce the processing speed. The current study result did not support it. The two conditions (H+F-P+ and H-F-P+) yielded identical response times. Another way to compare the same relationship is to hold homophone density constant and compare the two levels of tonal probability. Therefore, the results in Wiener and Ito (in press; $F-P+ < F-P-$) should predict a homophone effect, i.e. H+F-P+ should be faster than H+F-P-, and H-F-P+ should be faster than H-F-P-. The former relation ($H+F-P+ < H-F-P-$) was found in current study, but the latter relation ($H-F-P+ < H-F-P-$) was not shown.

The current study supported both *Independent Representation* and *Shared Representation models*. First, the homophones compete with each other during lexical processing. As the numerical difference between H+F+P+ and H-F+P+ suggests that low homophone density facilitated the processing and the facilitative effect is probability due to the less competition of homophones in the low homophone density condition as discussed above. Homophone mates with high specific frequency were activated earlier and slowed down the recognition of less frequent homophone words. On the other hand, the different values of H-F+P- and H+F+P- suggests that having homophone mates with higher frequency can facilitate the recognition and words with less highly frequent homophone mates will be responded more slowly. Briefly, tonal probability of words in high syllable frequency influenced whether homophone mates had a facilitative or inhibitory effect. In other words, the tonal probability influences the mental representation of homophones in high syllable frequency conditions.

A potential problem for current study is that the homophone density of non-target words in each slide, such as the homophone densities of segmental competitor, rhyme competitor, and distractor could not be controlled. Some differences between the current study and Wiener & Ito (in press) might be explained by the failure to control such factors. Wiener and Ito found no mistakes due to participants clicking on the rhyme competitor. However, in the current study, 22 errors were made by clicking on rhyme competitor, while the total number of mistaken clicks on segmental competitor was 25. In addition, the looking proportions to the distractor at the beginning stage were relatively high in some conditions, but they dropped quickly before 200ms from onset. Furthermore, in H+F-P- condition, there was a 'general confusion' state which is shown by fixation plots; that is, all the four lines were entangled together until the point where the target line had a clear rising. This phenomenon might be ascribed to the effect of homophone density. In the condition, the variation of homophone density for segmental competitor, rhyme competitor, and distractor might have had an effect on participants' looking.

In current study, only trends of homophone effect and tonal probability were shown by the analysis. Additional participants are currently being recruited to tease the three factors apart. In addition, researchers can also try to use other paradigms such as phonological priming, visual priming paradigm, or lexical decision tasks on pseudo-homophones to study the roles of homophones, syllable frequency, and tonal probability in Mandarin lexical processing. These paradigms have showed been successful in other languages such as English and Dutch.

Conclusion

In conclusion, the current study suggests a complex interaction of homophone density, syllable frequency, and tonal probability in lexical processing. As the results suggested, the automatic mental processing for infrequent syllables with improbable tone in a high homophone

density environment is most difficult, while it is easiest for frequent syllable with probable tone in low homophone density. The segmental frequency does not necessarily influence the processing speed, but did interact with homophone density. Homophone density delays the processing for infrequent syllable with improbable tone, but it facilitates the processing of frequent syllables with probable tone.

Acknowledgments

My deepest gratitude goes to my advisors Shari Speer and Kiwako Ito. Without their helps and guidance, none of this work could be done. Also, my thanks goes to Seth Wiener for his helps to in preparing experimental materials, recruiting participants, and assistance in statistical analysis.

Reference

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-41. Retrieved from: <http://CRAN.R-Project.org/package=lme4>

Biedermann, B., Blanken, G., & Nickels, L. (2002). The representation of homophones: Evidence from remediation. *Aphasiology*, 16, 1115-1136.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 27, 6, 1430-50.

Chen, C. Y., Tseng, S. F., Huang, C. R., & Chen, K. J. (1993). Some distributional properties of Mandarin Chinese—A study based on the Academia Sinica corpus. In *Proceedings of Pacific Asia Conference on Formal and Computational Linguistics I* (pp. 81-95).

Chen, H.-C., Vaid, J., & Wu, J.-T. (2009). Homophone density and phonological frequency in Chinese word recognition. *Language and Cognitive Processes*, 24, 967-982.

Duanmu, S., 1999. Stress and the development of disyllabic vocabulary in Chinese. *Diachronica XVI*: 1.1-35.

Duanmu, S., 2012. Syllable and syllable structure in Chinese. In *Encyclopedia of Chinese Language and Linguistics*, ed. Wolfgang Behr, Gu Yueguo, Zev Handel, C.-T. James Huang, & Rint Sybesma. Leiden: Brill

Hakekoh, U., & Højsgaard, S. (2015). Pbkrttest: Parametric bootstrap and Kenward-Roger-based methods for mixed model comparison. Retrived from: <http://cran.r-project.org/web/packages/pbkrttest/index.html>

Ke, J. (2006). A cross-linguistic quantitative study of homophony*. *Journal of Quantitative Linguistics*, 13(01), 129-159.

Kerswell, L., Siakaluk, P. D., Pexman, P. M., Sears, C. R., & Owen, W. J. (2007). Homophone effects in visual word recognition depend on homophone type and task demands. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Expérimentale*, 61, 4, 322-7.

Li, P., & Yip, M. C. (1996). Lexical ambiguity and context effects in spoken word recognition: Evidence from Chinese. In G. W. Cottrell (Ed.), *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society: July 12-15, 1996*. Mahwah, NJ: Lawrence Erlbaum Associates.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, 9(3), 225-236.

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 27, 1, 139-56.

Qiu, X. (2000). Chinese Writing. (Louis, G., & Norman, J. Trans.) In *Early China Special Monograph Series* (Vol. 4). Berkeley: Society for the Study of Early China and Institute of East Asian Studies, University of California, Berkeley.

Treiman, R., Seidenberg, M. S., & Kessler, B. (2014). Influences on spelling: evidence from homophones. *Language, Cognition and Neuroscience*, (ahead-of-print), 1-11.

Wen, J. (2012). *A statistical argument for the homophony avoidance approach to the disyllabification of Chinese*. In *Proceedings of the 23rd North American Conference on Chinese Linguistics (NACCL-23)*, ed. Zhou Jing-Schmidt, University of Oregon, Eugene.

Wiener, S., & Ito, K. (in press). Do syllable-specific tonal probabilities guide lexical access? Evidence from Mandarin, Shanghai and Cantonese speakers. *Language, Cognition and Neuroscience*, 1-13.

Wiener, S., Speer, S. R., & Shank, C. (2012). Effects of frequency, repetition and prosodic location on ambiguous Mandarin word production. In *Proceedings of the 6th International Conference on Speech Prosody* (pp. 528-531).

Appendix: Language Background Questionnaire (Translated)

Language Background Questionnaire

Age _____ Sex (M / F)

1. Where have you lived, and what age were you at the time (starting with place of birth)?
Place (City, Province) Ages (From.... To....)

2. How long have you been in the U.S.? _____

3. Where were your parents or other caretakers born?
Who? Place (City, Province)

4. Where were your grandparents born?
Who? Place (City, Province)

5. Did your parents or other caretakers speak any other languages other to you at home? If so, which one(s)? _____

6. What other languages and dialects do you speak (if any)? At what age did you begin learning these languages? How well can you read, write, speak and understand each one?

(1) barely / not at all (2) poorly (3) passably (4) fluently

Language/dialect:	Age:	Reading	Writing	Speaking	Listening
		1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
		1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
		1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
		1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

7. Are you married or have a domestic partner? If so, what language do you speak to your spouse/partner in?

8. What is the highest level of education you have completed in China?
primary junior high high school college post-graduate

9. What is your profession? (If academic or student, indicate field of study)

10. On average, how many minutes a day do you spend reading or writing in Chinese?

11. On average, what percentage of your daily conversations with others is in Chinese? For example, 50% in Chinese and 50% in English.

12. Do you have any speech or hearing disorders? Y / N (if yes, please describe)

13. Do you have any vision impairment? Y / N (if yes, please describe).

14. May we contact you for future studies? Y / N
If you answered yes, please provide your contact phone number and/or e-mail.
